Int J Game Theory (1998) 27:49–59

International Journal of Game Theory

Rationality and the definition of consistent pairs

Christian Ewerhart*

University of Bonn, Wirtschaftspolitische Abteilung, Adenauerallee 24-42, D-53113 Bonn, Germany

Received November 1997/First version May 1997

Abstract. A consistent pair specifies a set of "rational" strategies for both players such that a strategy is rational if and only if it is a best reply to a Bayesian belief that gives positive probability to every rational strategy of the opponent and probability zero otherwise. Although the idea underlying consistent pairs is quite intuitive, the original definition suffers from non-existence problems. In this article, we propose an alternative formalization of consistent pairs. According to our definition, a strategy is "rational" if and only if it is a best reply to some lexicographic probability system that satisfies certain consistency conditions. These conditions imply in particular that a player's probability system gives infinitely more weight to rational strategies than to other strategies. We show that modified consistent pairs exist for every game.

Key words: Consistent pairs, rationality, lexicographic probability systems

1. Introduction

The concept of consistent pairs developed by Börgers and Samuelson (1992) is motivated by the natural requirement that rational players maximize expected utility using "cautious" beliefs, i.e. beliefs that are formed with respect to the entire set of "rational" strategies. In the formal setting chosen by Bögers and

* This article is based on Chapter 3 of my Ph.D. thesis finished at the University of Bonn in fulfillment of the requirements of the European Doctoral Programme. For helpful comments and discussions, I would like to thank Eddie Dekel, Larry Blume, Tilman Börgers, Martin Dufwenberg, Frank Schuhmacher, Ariel Rubinstein, Avner Shaked, and seminar participants at Tel Aviv and Iowa City. Financial assistance by the German Academic Exchange Service is gratefully acknowledged.

Samuelson, rational players are Bayesian utility maximizers and cautious beliefs are defined as probability distributions that give positive probability to a strategy if and only if it is possibly played by a rational opponent. The approach results in a solution concept that yields attractive results for some games. Yet, it suffers from nonexistence problems in other games. Börgers and Samuelson conclude that it is logically inconsistent to assume that rational players analyze all games on the basis of cautious expected utility maximization.

In this article we propose a modified definition of consistent pairs using the non-Archimedian version of subjective expected utility theory developed by Blume, Brandenburger and Dekel (1991). Players will be assumed to adhere to this theory and cautious beliefs are formalized as lexicographic probability systems that satisfy certain consistency requirements. These requirements imply in particular that a cautious belief gives infinitely more weight to a "rational" strategy than to any other strategy. With this modification, we are able to prove existence for consistent pairs. To keep matters as simple as possible, we shall restrict ourselves to two-player games. The results of this article do not depend on this assumption (see Ewerhart (1997) for a more general treatment).

The rest of the article is organized as follows. In the next section we introduce the notation and recall the definition of lexicographic probability systems. In Section III, we present our definition of consistent pairs and discuss its properties. Section IV illustrates the theory by means of examples. Section V concludes. The Appendix contains the technical parts of the proof of the existence theorem.

II. Notations and definitions

A *finite two-person game in normal form* is a quadruple $\Gamma = \langle S_1, S_2; u_1, u_2 \rangle$ where for each $i = 1, 2, S_i$ is player i's finite set of pure strategies and $u_1 : S_1 \times S_2 \to \mathbb{R}$ is i's von Neumann-Morgenstern utility function. For any finite set Y, let $\Delta(Y)$ denote the set of all probability measures on Y. The set of *mixed strategies* of player i is then $\sum_i = \Delta(S_i)$, with a typical element denoted by σ_i. We call the set of pure strategies that have positive probability in a mixed strategy σ_i the *support* of σ_i and denote it by $\mathrm{Supp}(\sigma_i)$. If C_i is a set of mixed strategies then $\mathrm{conv}(C_i)$ denotes the *convex hull* of C_i. It represents the set of beliefs about C_i (cf. Pearce (1984), Lemmas 1 and 2 for details).

A strategy profile, i.e. a pair of mixed strategies will be written as $\sigma = (\sigma_1, \sigma_2) \in \sum_1 \times \sum_2$. The utility functions extend linearly to $\sum_1 \times \sum_2$. It will be helpful to use the formal symbol $-i$ to denote player i's opponent. A vector $\rho_i = (\sigma_i^1, \ldots, \sigma_i^{K_i})$ of probability distributions on S_i will be referred to as a *lexicographic probability system* (LPS). The first component σ_i^1 of the LPS ρ_i can be thought of as representing the decision maker's primary belief about the strategy choice of player i, the second component σ_i^2 as representing the decision maker's secondary belief, and so on. The set of best responses to an LPS is defined as follows. For the "empty" probability system $\rho_{-i} = (\ )$, let $\mathrm{BR}_i\ (\rho_{-i}) = S_i$. Then, by induction, let

$$\mathrm{BR}_i(\sigma_{-i}^1, \ldots, \sigma_{-i}^k) = \{s_i \in \mathrm{BR}_i(\sigma_{-i}^1, \ldots, \sigma_{-i}^{k-1}) | u_i(s_i, \sigma_{-i}^k) \geq u_i(s_i', \sigma_{-i}^k) \text{ for all } s_i' \in \mathrm{BR}_i(\sigma_{-i}^1, \ldots, \sigma_{-i}^{k-1})\}. \tag{1}$$

Given some lexicographic probability system $\rho_{-i} = (\sigma_{-i}^1, \dots, \sigma_{-i}^{K_i})$ of player $-i$ on the strategy of player i, this defines $\mathrm{BR}_i(\rho_{-i})$, i.e. the *set of pure best replies* to ρ_{-i}.

III. Consistent pairs

The set of "rational" strategies for a player certainly depends on what is rational for the other player. We shall therefore investigate the implications of imposing the following kind of axiom.

> For each player, a strategy is rational if and only if it is a best reply, given the set of rational strategies of the other player.

In this section we shall give this intuitive requirement a formal setting.

Define the *support of a set* $C_i \subseteq \sum_i$ for player i in the obvious way as $\mathrm{Supp}(C_i) = \bigcup_{\sigma_i \in C_i} \mathrm{Supp}(\sigma_i)$, i.e. the union of the supports of all mixed strategies in C_i. By a *collection of theories* for player i we mean an L_i-tuple $\mathscr{C}_i = (C_i^1, \dots, C_i^{L_i})$ of sets $C_i^l \subseteq \sum_i$ such that $\mathrm{Supp}(C_i^l) \cap \mathrm{Supp}(C_i^{l'}) = \phi$ if $l \neq l'$ (the requirement on the supports can be disposed of by minor modifications of the subsequent definitions, cf. Ewerhart (1997)). Similar to the interpretation of an LPS, the first component C_i^1 of a collection of theories can be interpreted as the primary theory of the decision maker about the strategy choice of player i, the second component C_i^2 as the secondary theory, and so on. A collection of theories $\mathscr{C}_i = (C_i^1, \dots, C_i^{L_i})$ will be called *complete* if for every $s_i \in S_i$ there exists an $l \in \{1, \dots, L_i\}$ such that $s_i \in \mathrm{Supp}(C_i^l)$. We shall say that a lexicographic probability system $\rho_i = (\sigma_i^1, \dots, \sigma_i^{K_i})$ is *consistent with the collection of theories* $\mathscr{C}_i = (C_i^1, \dots, C_i^{L_i})$ if there are indices $0 = k_0 \leq k_1 \leq \cdots \leq k_L = K_i$, such that for all $l \in \{1, \dots, L_i\}$ the following two conditions are satisfied:

1. for all $k \in \{1, \dots, K_i\}$, if $k_{l-1} < k \leq k_l$ then $\sigma_i^k \in \mathrm{conv}(C_i^l)$, and
2. $\mathrm{Supp}(C_i^l) = \bigcup_{k=k_{l-1}+1}^{k_l} \mathrm{Supp}(\sigma_i^k)$.

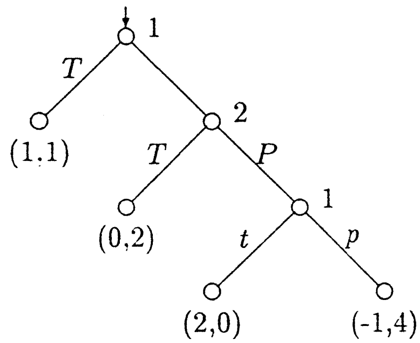
The intuition behind this definition is the following. If some collection of theories $\mathscr{C}_i = (C_i^1, \dots, C_i^{L_i})$ applies then only those beliefs systems ρ_i are considered to be "rational" which are consistent with $\mathscr{C}_i$. In the main part, this implies that every pure strategy in $\mathrm{Supp}(C_i^l)$ is expected to be infinitely more likely than any strategy in $\mathrm{Supp}(C_i^{l+1})$.

For the definition of modified consistent pairs, we shall need some more notation. Fix some collection of theories $\mathscr{C}_{-i}$ for player $-i$. Denote by $\mathrm{MBR}_i(\mathscr{C}_{-i})$ the set of all $\sigma_i \in \sum_i$ such that there exists an LPS ρ_{-i} consistent with $\mathscr{C}_{-i}$ such that σ_i gives positive weight only to strategies in $\mathrm{BR}_i(\rho_{-i})$. The set $\mathrm{MBR}_i(\mathscr{C}_{-i})$ will be called the *set of mixed best replies to the collection of theories* $\mathscr{C}_{-i}$. Finally, if $\mathscr{C}_i = (C_i^1, \dots, C_i^{L_i})$ is a collection of theories for player i, and $A_i \subseteq \sum_i$, then $(A_i, \mathscr{C}_i)$ denotes the collection of theories $(A_i, C_i^1, \dots, C_i^{L_i})$.

Definition 1. *Given a set $A_i \subseteq \sum_i$ for every player $i = 1, 2$, the pair (A_1, A_2) is called a* **modified consistent pair** *if there exist collections of theories $\mathscr{C}_1$ and $\mathscr{C}_2$ such that*

1. $(A_i, \mathscr{C}_i)$ *is a complete collection of theories for* $i = 1, 2$, *and*
2. $A_i = \text{MBR}_i(A_{-i}, \mathscr{C}_{-i})$ *for* $i = 1, 2$.

In this case, the collection of theories $\mathscr{C}_1, \mathscr{C}_2$ *are said to* support *the pair* (A_1, A_2).



	T	P
T	1,1	1,1
t	0,2	2,0
p	0,2	-1,4

Fig. 1

To illustrate the definition consider a simple centipede game (Figure 1). In this game, (T, T) is a modified consistent pair, where we leave away the parentheses for the sake of notational simplicity. To support this pair, choose collections $\mathscr{C}_1 = (\{t\}, \{p\})$ and $\mathscr{C}_2 = (\{P\})$. The interpretation is the following. Player 1 chooses T as be believes that Player 2 plays T. Player 2 believes first order that Player 1 plays T. His decision will be forced by his second order belief. Since he estimates t to be infinitely more likely than p, Player 2 plays T.

Theorem 1. *For any two-person game, there exists a modified consistent pair.*

Proof: We define inductively, simultaneously for players 1 and 2, a sequence $(A_i^{(n)}, \mathscr{C}_i^{(n)})_{n \geq 1}$ where for each n, $A_i^{(n)} \subseteq \sum_i$ and $\mathscr{C}_i^n$ is a collection of theories for player i. For this, let $A_i^{(1)} = \sum_i$ and $\mathscr{C}_i^{(1)} = (\ )$. Now, for a given $n > 1$, define

$$A_i^{(n)} = \text{MBR}(A_{-i}^{(n-1)}, \mathscr{C}_{-i}^{(n-1)}) \tag{2}$$

and

$$\mathscr{C}_{-i}^{(n)} = \begin{cases} \mathscr{C}_{-i}^{(n-1)} & \text{if } \text{Supp}(A_i^{(n)}) = \text{Supp}(A_i^{(n-1)}), \\ (S_i \cap (A_i^{(n-1)} \backslash A_i^{(n)}), \mathscr{C}_{-i}^{(n-1)}) & \text{otherwise.} \end{cases} \tag{3}$$

The following lemma concludes the proof of the theorem.

Lemma 1. *The above defined sequence* $(A_i^{(n)}, \mathscr{C}_i^{(n)})_{n \geq 1}$ *becomes stationary after a finite number of steps, yielding a modified consistent pair* $(A_1^{\infty}, A_2^{\infty})$ *and two collections of theories* $\mathscr{C}_1^{\infty}$ *and* $\mathscr{C}_2^{\infty}$ *that support it.*

The proof of Lemma 1 is postponed to the Appendix. q.e.d.

The induction process described by (2) and (3) determines the modified consistent pair (T, T) of the game in Figure 1. Following the the inductive definition step by step, we get:

$$
\begin{aligned}
&A_1^{(1)} = \sum\nolimits_1 && A_2^{(1)} = \sum\nolimits_2 \\
&\mathscr{C}_1^{(1)} = (\ ) && \mathscr{C}_2^{(1)} = (\ ) \\
&A_1^{(2)} = \Delta\{T, t\} && A_2^{(2)} = \sum\nolimits_2 \\
&\mathscr{C}_1^{(1)} = (\{p\}) && \mathscr{C}_2^{(1)} = (\ ) \\
&A_1^{(3)} = \Delta\{T, t\} && A_2^{(3)} = \{T\} \\
&\mathscr{C}_1^{(1)} = (\{p\}) && \mathscr{C}_2^{(1)} = (\{P\}) \\
&A_1^{(4)} = \{T\} && A_2^{(4)} = \{T\} \\
&\mathscr{C}_1^{(1)} = (\{t\}, \{p\}) && \mathscr{C}_2^{(1)} = (\{P\})
\end{aligned}
\tag{4}
$$

For $n \geq 5$, the sequence becomes stationary, i.e. $A_i^{(n)} = A_i^{(n-1)}$ and $\mathscr{C}_i^{(n)} = \mathscr{C}_i^{(n-1)}$. The respective limits are $A_1^{\infty} = \{T\}, A_2^{\infty} = \{T\}, \mathscr{C}_1^{\infty} = (\{t\}, \{p\})$ and $\mathscr{C}_2^{\infty} = (\{P\})$. This is the solution we described above.

In the remaining part of this section, we state and prove some of the properties of modified consistent pairs. First we make a technical remark. By a *subsimplex* of $\sum_i = \Delta(S_i)$, we mean a set of the form $\Delta(S_i')$ with some set $S_i' \subseteq S_i$.

Lemma 2. *For every collection of theories $\mathscr{C}_{-i}$, the set $\mathrm{MBR}_i(\mathscr{C}_{-i})$ of mixed best replies to $\mathscr{C}_{-i}$ is a union of subsimplices of $\sum_i$.*

Proof: For every lexicographic belief system ρ_{-i} that is consistent with $\mathscr{C}_{-i}$ the set $\mathrm{MBR}_i(\rho - i)$ is a subsimplex of $\sum_i$ by definition. As $\mathrm{MBR}_i(\mathscr{C}_{-i})$ is a union of sets $\mathrm{MBR}_i(\rho - i)$ the lemma follows. q.e.d.

Note that Lemma 2 implies that there exist at most finitely many modified consistent pairs for a given game. We proceed by discussing in what sense modified consistent pairs model cautious behaviour. Fix a set of strategies $B_{-i} \subseteq \sum_{-i}$ and call a strategy $\sigma_i \in \sum_i$ *weakly dominated against* $B_{-i} \subseteq \sum_{-i}$ if there exists a $\sigma_i' \in \sum_i$ such that $u_i(\sigma_i', \sigma_{-i}) \geq u_i(\sigma_i, \sigma_{-i})$ for every $\sigma_{-i} \in B_{-i}$ and the inequality is strict for at least one $\sigma_{-i} \in B_{-i}$. A "cautious" player would not use a strategy that is weakly dominated against the whole strategy set of the opponent. Similary, he would not use a strategy that is weakly dominated against the set of "rational" strategies of the opponent. We can show that our solution concept satisfies both requirements.

Theorem 2. *Let (A_1, A_2) be a modified consistent pair and let $\sigma_i \in A_i$ for some i. Then σ_i is neither weakly dominated against $\sum_{-i}$ nor weakly dominated against A_{-i}.*

Proof: Almost immediate from Proposition 1 in Brandenburger (1991). The only thing to show is that a strategy $\sigma_i \in \sum_i$ that is weakly dominated against $\mathrm{Supp}(A_{-i})$ is weakly dominated against A_{-i} itself. But this is obvious, since A_{-i} contains its support by Lemma 2. q.e.d.

We close this section with two simple results that illustrate the relation to standard solution concepts. For a more complete analysis along the same lines we refer the reader to Ewerhart (1997), Section 3.V.

Proposition 1. *If* (σ_1, σ_2) *is a strict Nash equilibrium then* $(\{\sigma_1\}, \{\sigma_2\})$ *is a modified consistent pair.*

Proof: As (σ_1, σ_2) is a strict Nash equilibrium, σ_i is a pure strategy for $i = 1, 2$. Let $\mathscr{C}_i = (S_i \backslash \{\sigma_i\})$, $i = 1, 2$. Then it is immediate that $\mathscr{C}_1$ and $\mathscr{C}_2$ support the pair $(\{\sigma_1\}, \{\sigma_2\})$. q.e.d.

Proposition 2. *Every strategy contained in a modified consistent pair is normal-form rationalizable in the sense of Bernheim (1984) and Pearce (1984).*

Proof: Let (A_1, A_2) be a modified consistent pair. Then each strategy σ_i in A_i in particular is a best reply to some belief about $A_{-i}, i = 1, 2$. Iterating this argument infinitely often yields a chain of justification that shows that σ_i is rationalizable. q.e.d.

IV. Examples

Proposition 3. *In the centipede game, the unique consistent pair is* (T, T).

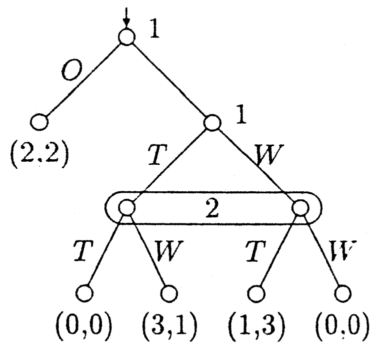
Proof: Let (A_1, A_2) be some consistent pair supported by collections of theories $\mathscr{C}_1$ and $\mathscr{C}_2$. Then we have $A_i = \mathrm{MBR}_i(A_{-i}, \mathscr{C}_{-i})$ for $i = 1, 2$. By Lemma 2, there are only three candidates for A_2, which will be treated separately.

1. Suppose $A_2 = \{P\}$. Since the unique best reply to $\{P\}$ is t, we may infer $A_1 = \{t\}$, implying that $P \notin A_2$, a contradiction. Therefore, $A_2 \neq \{P\}$.
2. Similarly, if $A_2 = \sum_2$ then $A_1 = \Delta\{T, t\}$, hence $P \notin A_2$, leading again a contradiction. This shows $A_2 \neq \sum_2$.
3. Finally, if $A_2 = \{T\}$ then necessarily $A_1 = \{T\}$, rendering the modified consistent pair (T, T).

Since all cases have been covered, the proof is complete. q.e.d.

Consider now the game in Figure 2, which is a variation of a game of Kohlberg and Mertens (1986).

We determine the set of modified consistent pairs for this game. Since the profile (T, W) is a strict Nash equilibrium one can infer from Proposition 1 that (T, W) is a consistent pair in this game. It is supported by any theory of irrational behavior, e.g. by $\mathscr{C}_1 = (\{O, W\})$ and $\mathscr{C}_2 = (\{T\})$. There is another consistent pair in this game, viz. (O, T). To support the pair (O, T), a collection of theories $\mathscr{C}_1$ must give W a higher probability rank than T. This requirement is satisfied e.g. by $\mathscr{C}_1 = (\{W\}, \{T\})$ and $\mathscr{C}_2 = (\{W\})$.



	T	W
O	2,2	2,2
T	0,0	3,1
W	1,3	0,0

Fig. 2

Proposition 4. *For the game in Figure 2, there exist exactly two modified consistent pairs.*

Proof: Let (A_1, A_2) be a modified consistent pair. By Lemma 2, there are three candidates for the set A_2. Firstly, let $A_2 = \{W\}$. As T is the unique best reply of Player 1 to W, necessarily $A_1 = \{T\}$, leading to the first of the two consistent pairs described above. Similarly, if $A_2 = \{T\}$, then $A_1 = \{O\}$, giving back the second consistent pair. Finally, assume that $A_2 = \Delta\{T, W\}$. Then $A_1 = \text{MBR}_1(A_2) = \Delta\{O, T\}$. But for Player 2, T is weakly dominated against A_1, hence by Theorem 2 not in A_2, a contradiction. As the are no other choices for A_2 the proof is complete. q.e.d.

Proposition 4 shows that non-uniqueness may lead to a modified consistent pair that ignores the forward-induction logic. We close with a perfect-information game for which one of two existing pairs violates backwards induction.

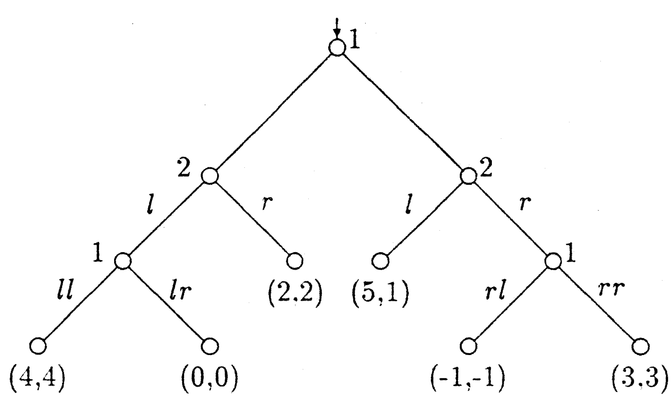


Fig. 3

In the game depicted in Figure 3, Player 2 has four pure strategies, where by way of example *rl* denotes the strategy that uses action *r* at the left decision node and action l at the right one. We determine the set of modified consistent pairs for this game.

Proposition 5. *In the game of Figure 3, there are exactly two modified consistent pairs, namely* (ll, lr) *and* (rr, rr).

Proof: Suppose, (A_1, A_2) is a consistent pair supported by collections of theories $\mathscr{C}_1$ and $\mathscr{C}_2$. Then, since lr and rl are weakly dominated for Player 1, they are not elements of A_1 by Theorem 2. Therefore, by means of Lemma 2, we know that A_1 is a subsimplex of $\Delta\{ll, rr\}$. We distinguish three cases.

1. Let $A_1 = \{ll\}$. Then Player 2 optimally replies by some element in $\Delta\{ll, lr\}$. But $ll \notin A_2$ because otherwise ll could not be the unique element of A_1. By Lemma 2, we may infer $A_2 = \{lr\}$. Now it is straightforward to check that (ll, lr) in fact is a modified consistent pair, supported e.g. by $\mathscr{C}_1 = (\{rr\}, \{lr, rl\})$ and $\mathscr{C}_2 = (\{ll, rl, rr\})$.
2. Assume now that $A_2 = \{rr\}$. Then $A_1 \subseteq \mathrm{MBR}(A_2) = \Delta\{lr, rr\}$. However, $lr \notin A_2$ because otherwise $ll \in A_1$. Again by Lemma 2, we get $A_2 = \{rr\}$ and one can see that this defines a modified consistent pair, which is supported e.g. by $\mathscr{C}_1 = (\{lr\}, \{ll, rl\})$ and $\mathscr{C}_2 = (\{ll, rl, rr\})$.
3. In the remaining case, viz. $A_1 = \Delta\{ll, rr\}$, we see that $A_2 = \{lr\}$ is the unique best reply. This leads to $rr \notin A_1$, a contradiction.

This proves the assertion. q.e.d.

V. Conclusion

A new definition of rationality in two person normal form games is proposed. The solution concept captures some of the intuition that underlies the idea of cautious utility maximization. Furthermore, although the concept is defined by a seemingly strong axiomatic requirement, it exists in any game. However, there are games for which our definition of rationality does not lead to unique predictions.

Appendix

In this appendix we prove Lemma 1, which is central for the proof of the existence result. We shall need the following lemma.

Lemma 3. *Let* $A_i \subseteq \sum_i$ *and* $A_i', A_i'' \subseteq A_i$ *such that*

$$\mathrm{Supp}(A_i') \cap \mathrm{Supp}(A_i'') = \phi, \tag{5}$$

$$\mathrm{Supp}(A_i') \cup \mathrm{Supp}(A_i'') = \mathrm{Supp}(A_i), \tag{6}$$

and let C_i *be some collection of theories for player i. Then*

$$\mathrm{MBR}_{-i}(A_i', A_i'', \mathscr{C}_i) \subseteq \mathrm{MBR}_{-i}(A_i, \mathscr{C}_i). \tag{7}$$

Proof: We show that every lexicographic probability system that is consistent with the sequence $(A_i', A_i'', \mathscr{C}_i)$ is consistent with $(A_i, \mathscr{C}_i)$ as well. Assume that $\rho_i = (\sigma_i^1, \ldots, \sigma_i^K)$ is consistent with $(A_i', A_i'', C_i^1, \ldots, C_i^L)$. Then there is a finite sequence $0 \leq k' \leq k_0 \leq k_1 \leq \cdots \leq k_L = K$ such that

$$\forall k \in \{1, \ldots, k'\} : \sigma_i^k \in \operatorname{conv}(A_i'), \tag{8}$$

$$\forall k \in \{k'+1, \ldots, k_0\} : \sigma_i^k \in \operatorname{conv}(A_i''), \tag{9}$$

$$\bigcup_{k \in \{1, \ldots, k'\}} \operatorname{Supp}(\sigma_i^k) = \operatorname{Supp}(A_i'), \tag{10}$$

$$\bigcup_{k \in \{k'+1, \ldots, k_0\}} \operatorname{Supp}(\sigma_i^k) = \operatorname{Supp}(A_i''), \tag{11}$$

and for every $l \in \{1, \ldots, L\}$,

$$\forall k \in \{k_{l-1}+1, \ldots, k_l\} : \sigma_i^k \in \operatorname{conv}(C_i^l), \tag{12}$$

$$\bigcup_{k \in \{k_{l-1}+1, \ldots, k_l\}} \operatorname{Supp}(\sigma_i^k) = \operatorname{Supp}(C_i^l). \tag{13}$$

Since $\operatorname{conv}(A_i') \cup \operatorname{conv}(A_i'') \subseteq \operatorname{conv}(A_i)$ and $\operatorname{Supp}(A_i') \cup \operatorname{Supp}(A_i'') = \operatorname{Supp}(A_i)$ by assumption, there exist $0 \leq k_0 \leq k_1 \leq \cdots \leq k_L = K$ such that

$$\forall k \in \{1, \ldots, k_0\} : \sigma_i^k \in \operatorname{conv}(A_i), \tag{14}$$

$$\bigcup_{k \in \{1, \ldots, k_0\}} \operatorname{Supp}(\sigma_i^k) = \operatorname{Supp}(A_i), \tag{15}$$

and for every $l \in \{1, \ldots, L\}$, equations (12) and (13) are satisfied. Thus, ρ_i is consistent with $(A_i, \mathscr{C}_i)$. q.e.d.

Lemma 4. *The sequence $A_i^{(n)}$ is nested, i.e. $A_i^{(n+1)} \subseteq A_i^{(n)}$ for $n \geq 1, i = 1, 2$. Furthermore, $(A_i^{(n)}, \mathscr{C}_i^{(n)})$ is a complete collection of theories for all $n \geq 1$ and $i = 1, 2$.*

Proof: By induction with respect to n. For $n = 1$, the assertion follows immediately from $A_i^{(1)} = \sum_i$. Suppose now that, for some $n > 1, A_i^{(n)} \subseteq A_i^{(n-1)}$ and $(A_i^{(n)}, C_i^{(n)})$ is a complete collection of theories for $i = 1, 2$. We distinguish two cases.

1. Let $\operatorname{Supp}(A_{-i}^{(n)}) = \operatorname{Supp}(A_{-i}^{(n-1)})$. Then the definition of the sequence gives $\mathscr{C}_i^{(n)} = \mathscr{C}_i^{(n-1)}$. Therefore,

$$A_i^{(n+1)} = \mathrm{MBR}_i(A_{-i}^{(n)}, \mathscr{C}_{-i}^{(n)}) \tag{16}$$

$$\subseteq \mathrm{MBR}_i(A_{-i}^{(n-1)}, \mathscr{C}_{-i}^{(n-1)}) \tag{17}$$

$$= A_i^{(n)}, \tag{18}$$

where the inclusion follows from the induction hypothesis.

2. Now suppose that $\mathrm{Supp}(A_{-i}^{(n)}) \neq \mathrm{Supp}(A_{-i}^{(n-1)})$. Then $\mathscr{C}_i^{(n)} = (\tilde{A}_{-i}^{(n)}, \mathscr{C}_{-i}^{(n-1)})$, where $\tilde{A}_{-i}^{(n)} = S_{-i} \cap (A_{-i}^{(n-1)} \backslash A_{-i}^{(n)})$. We assume for a moment that the conditions of Lemma 3 are satisfied. Then

$$A_i^{(n+1)} = \mathrm{MBR}_i(A_{-i}^{(n)}, \mathscr{C}_{-i}^{(n)}) \tag{19}$$

$$= \mathrm{MBR}_i(A_{-i}^{(n)}, \tilde{A}_{-i}^{(n)}, \mathscr{C}_{-i}^{(n-1)}) \tag{20}$$

$$\subseteq \mathrm{MBR}_i(A_{-i}^{(n-1)}, \mathscr{C}_{-i}^{(n-1)}) \tag{21}$$

$$= A_i^{(n)}, \tag{22}$$

where the first equality follows from $\mathscr{C}_{-i}^{(n)} = \tilde{A}_{-i}^{(n)}, \mathscr{C}_{-i}^{(n-1)})$ and the inclusion from Lemma 3. It remains to check the conditions of Lemma 3. Firstly, by the induction hypothesis, we have $A_{-i}^{(n)} \subseteq A_{-i}^{(n-1)}$. Secondly, $\tilde{A}_{-i}^{(n)} \subseteq A_i^{(n-1)}$ by the definition of $\tilde{A}_{-i}^{(n)}$. Thirdly, we must show that $\mathrm{Supp}(A_{-i}^{(n)}) \cap \mathrm{Supp}(\tilde{A}_{-i}^{(n)}) = \phi$. For this, take some $s_{-i} \in \mathrm{Supp}(A_{-i}^{(n)})$. Then, by Lemma 2, in fact $s_{-i} \in A_{-i}^{(n)}$, implying that $s_{-i} \notin \tilde{A}_{-i}^{(n)}$. Since $\tilde{A}_{-i}^{(n)}$ consists of pure strategies only, $s_{-i} \in \mathrm{Supp}(\tilde{A}_{-i}^{(n)})$, showing that the supports of $A_{-i}^{(n)}$ and $\tilde{A}_{-i}^{(n)}$ are disjoint. Finally, we have to prove that $\mathrm{Supp}(A_{-i}^{(n)}) \cup \mathrm{Supp}(\tilde{A}_{-i}^{(n)}) = \mathrm{Supp}(A_{-i}^{(n-1)})$. Suppose that $s_{-i} \in \mathrm{Supp}(A_{-i}^{(n-1)}) \backslash \mathrm{Supp}(A_{-i}^{(n)})$. Then, in particular, $s_{-i} \notin A_{-i}^{(n)}$ and, by Lemma 2, $s_{-i} \in A_{-i}^{(n-1)}$, leading to $s_{-i} \in \tilde{A}_{-i}^{(n)}$ and hereby proving the assertion concerning the supports.

This closes the induction argument and proves Lemma 4. q.e.d.

Proof of Lemma 1: For $i = 1, 2$, the sequence $(A_i^{(n)})_{n \geq 1}$ is nested by Lemma 4. Moreover, there are only finitely many possible values for elements of this sequence by Lemma 2. Therefore, there exist an integer $n_0 \geq 1$ and sets $A_i^{\infty} \subseteq \sum_i$ for $i = 1, 2$ such that $A_i^{\infty} = A_i^{(n)}$ for all $n \geq n_0$. From the definition of the $\mathscr{C}_i^{(n)}$ we may infer that there exist a collection of theories $\mathscr{C}_i^{\infty}$ such that $\mathscr{C}_i^{(n)} = \mathscr{C}_i^{\infty}$ for all $n \geq n_0$ and $i = 1, 2$. Finally, the definition of $A_i^{(n)}$ gives $A_i^{\infty} = \mathrm{MBR}_i(A_{-i}^{\infty}, \mathscr{C}_{-i}^{\infty})$ for $i = 1, 2$. Since Lemma 4 ensures that $(A_i^{\infty}, \mathscr{C}_i^{\infty})$ is a complete theory for $i = 1, 2$, the proof of Lemma 1 is complete. q.e.d.

References

[1] Bernheim BD (1984) Rationalizable strategic behavior. Econometrica 52:1007–1028

[2] Blume L, Brandenburger A, Dekel E (1991) Lexicographic probabilities and choice under uncertainty. Econometrica 59:61–79

[3] Börgers T, Samuelson L (1992) 'Cautious' utility maximization and iterated weak dominance. International Journal of Game Theory 21:13–25

[4] Brandenburger A (1991) Lexicographic probabilities and iterated admissibility. In: Gale D, Hart O (ed.) Economic analysis of markets and games, MIT Press

[5] Ewerhart C (1997) On strategic reasoning and theories of rational behavior. Ph.D. thesis, University of Bonn
[6] Kohlberg E, Mertens JF (1986) On the strategic stability of equilibria. Econometrica 54:1003–1037
[7] Pearce DG (1984) Rationalizable strategic behavior and the problem of perfection. Econometrica 52:1029–1050